USGS
science for a changing world

Open Data and the USGS Science Data Catalog

PNAMP/CDI DMWG Seminar Series
February 17, 2016

U.S. Department of the Interior
U.S. Geological Survey

Goals

1. Complete the Science Data Lifecycle!!!!
2. What is Open Data?
3. U.S. gov actions, requirements for Open Data
4. Implementation at USGS and beyond
    a. Overview, history of USGS Science Data Catalog (SDC), pathways to data.doi.gov, data.gov, and the Geospatial Platform
    b. Revamped Science Data Catalog, late 2015
5. Future plans and goals related to Open Data

USGS

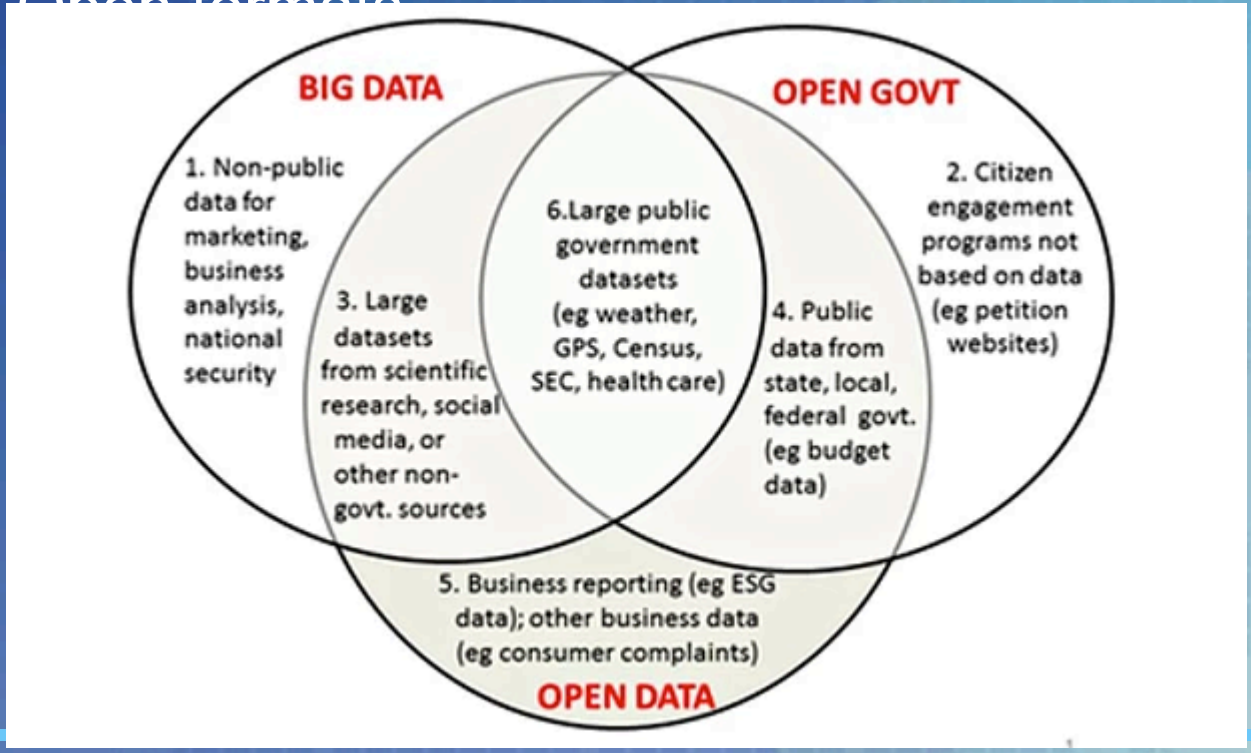
What is Open Data?

* How does it differ from “regular data”?
  * Open formats
  * Licenses
  * Collaboration
  * Machine interaction
  * Semanticyness
  * Branding

USGS

What is Open Data?

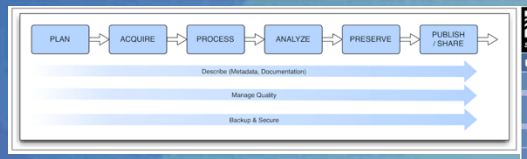
* How does it differ from “regular data”?
  * Open formats
  *
  *
  *
  *
  *
  *



USGS

Background & Drivers…

1. On May 9, 2013, the White House released the Executive Order, “Making Open and Machine Readable the New Default for Government.” This Order, built upon an earlier interagency memorandum released by the OSTP, was accompanied by an OMB Policy requiring:
   - Inventory of Data
   - Public Data Listing
     - data.doi.gov
     - data.gov
   - Project Open Data: https://project-open-data.cio.gov/
2. USGS Science Data Management activities
   - Data Management Instructional Memoranda
   - Community for Data Integration
   - Moving toward a “science data life cycle approach” in USGS
   - Implementation of USGS Science Strategy Plans



USGS

Open Government

- The Open Government Directive of December 2009 focuses on ***transparency, participation,*** and ***collab***
- Objectives:
  - Publish government information online
  - Improve the quality of information
  - Education and culture alignment
  - Create a policy framework for Open Government

USGS

The Digital Strategy

- On May 23, 2012 the Administration published "Digital Government: Building a 21st Century Platform to Better Serve the American People"
- Built on four guiding principles:
  - Information-Centric
  - Shared Platform
  - Customer-Centric
  - Security and Privacy

USGS

Digital Gov. Strategy Principles

1. An "***Information-Centric"*** approach—Moves us from managing "documents" to managing discrete pieces of open data and content which can be tagged, shared, secured, mashed up and presented in the way that is most useful for the consumer of that information.
2. A ***"Shared Platform"*** approach—Helps us work together, both within and across agencies, to reduce costs, streamline development, apply consistent standards, and ensure consistency in how we create and deliver information.
3. A ***"Customer-Centric"*** approach—Influences how we create, manage, and present data through websites, mobile applications, raw data sets, and other modes of delivery, and allows customers to shape, share and consume information, whenever and however they want it.
4. A platform of ***"Security and Privacy"***—Ensures this innovation happens in a way that ensures the safe and secure delivery and use of digital services to protect information and privacy.

USGS

Drivers (cont…)

* Open Data Action Plan
* Inclusion in quarterly “Cross Agency Priority” (CAP) goals
* Interagency Working Groups
* Policy Teams
* FOIA Requests
* Lobbying
* Etc….

USGS

Focused Open Data Initiatives

- "Button" projects
  - Green Button--energy usage, http://www.greenbuttondata.org/
  - Blue Button--health records, https://www.healthit.gov/patients-families/your-health-data
- Focused sites:
  - College matching/data: https://collegescorecard.ed.gov/
  - National Climate Assessment: http://www.globalchange.gov/browse/datasets
- Data.gov Communities
- Geospatial Platform Communities
  - National Geospatial Data Assets (NGDA)

USGS

Engaging the Public

- Data Jam
  - A closed-press, day-long ideation event with developers, designers, and subject matter experts focused on one topic and top related open data sets. Several are held in succession, leading up to a datapalooza three months later. Ex: Health Data jam (HHS), 21st Century Jobs Jam (OVP, Commerce, OSTP), Mitigating Campus Sexual Assault (Department of Education, Department of Justice)
- Datapalooza
  - An open press celebration, demo day, and platform to announce government open data releases or improvements. Ex: Safety Datapalooza (DOT, CPSC, FDA.)
- Hackathon
  - An event where developers, designers, and strategists work in teams to solve problems with software and/or hardware and demo the resulting work at the end of the day.Ex: White House "We The People" API Hackathon, The American Art API Hackathon
- "Sunshine Week": http://sunshineweek.rcfp.org/ et al.
- Challenges
  - "App-lifying USGS Data"
  - Climate Resilience Data
  - etc. https://www.data.gov/climate/climate-challenges

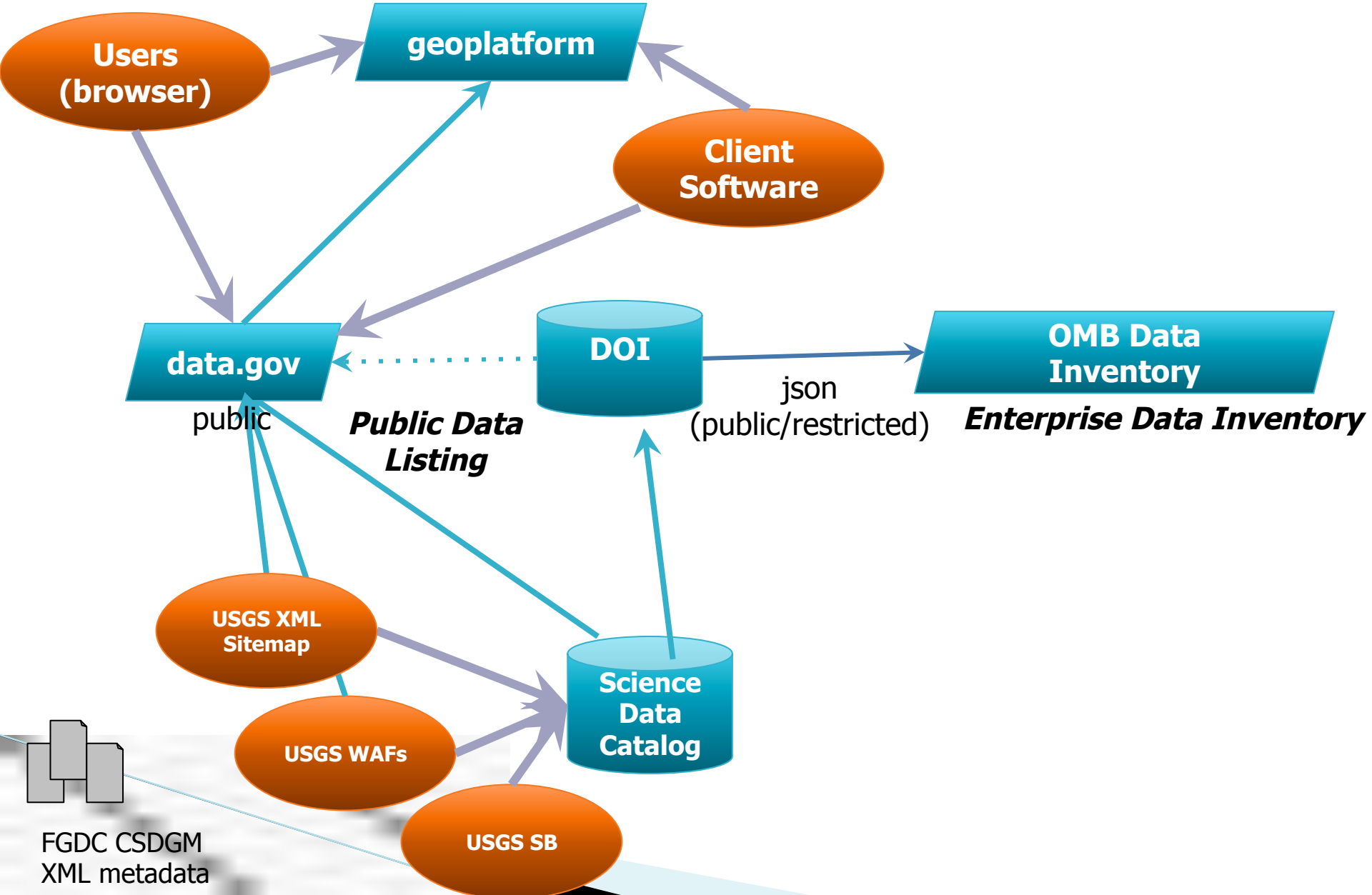
USGS

What is the USGS Science Data Catalog (SDC)?

1. Single point for USGS to provide its metadata to data.doi.gov, data.gov, OMB, geoplatform etc.
2. Organizes USGS data
3. Meets Open Data requirements for USGS
4. Provides metadata retrieval, visualization, download, & links back to original providers
5. Focus on scientific data

USGS

Data Publishing for Open Data and data.gov

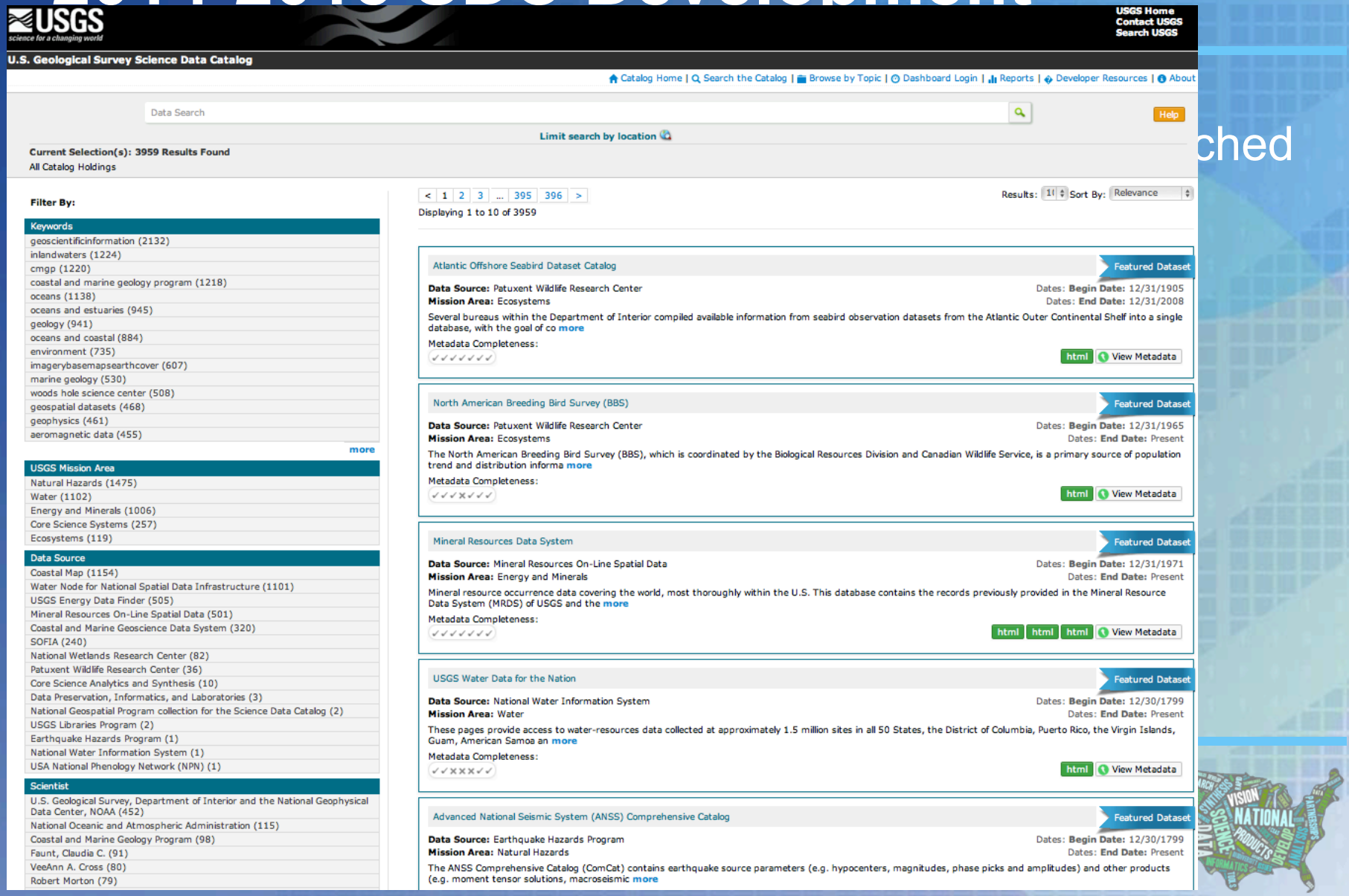


2014-2015 SDC Development

- Original version of SDC developed late 2013, launched early 2014
  - ~3500 initial records
- Integrated new data sources from several centers
- Priority datasets:
  - National Geospatial Data Asset (NGDA)
  - Big Earth Data Initiative (BEDI)
  - Data.gov communities

USGS

2014-2015 SDC Development



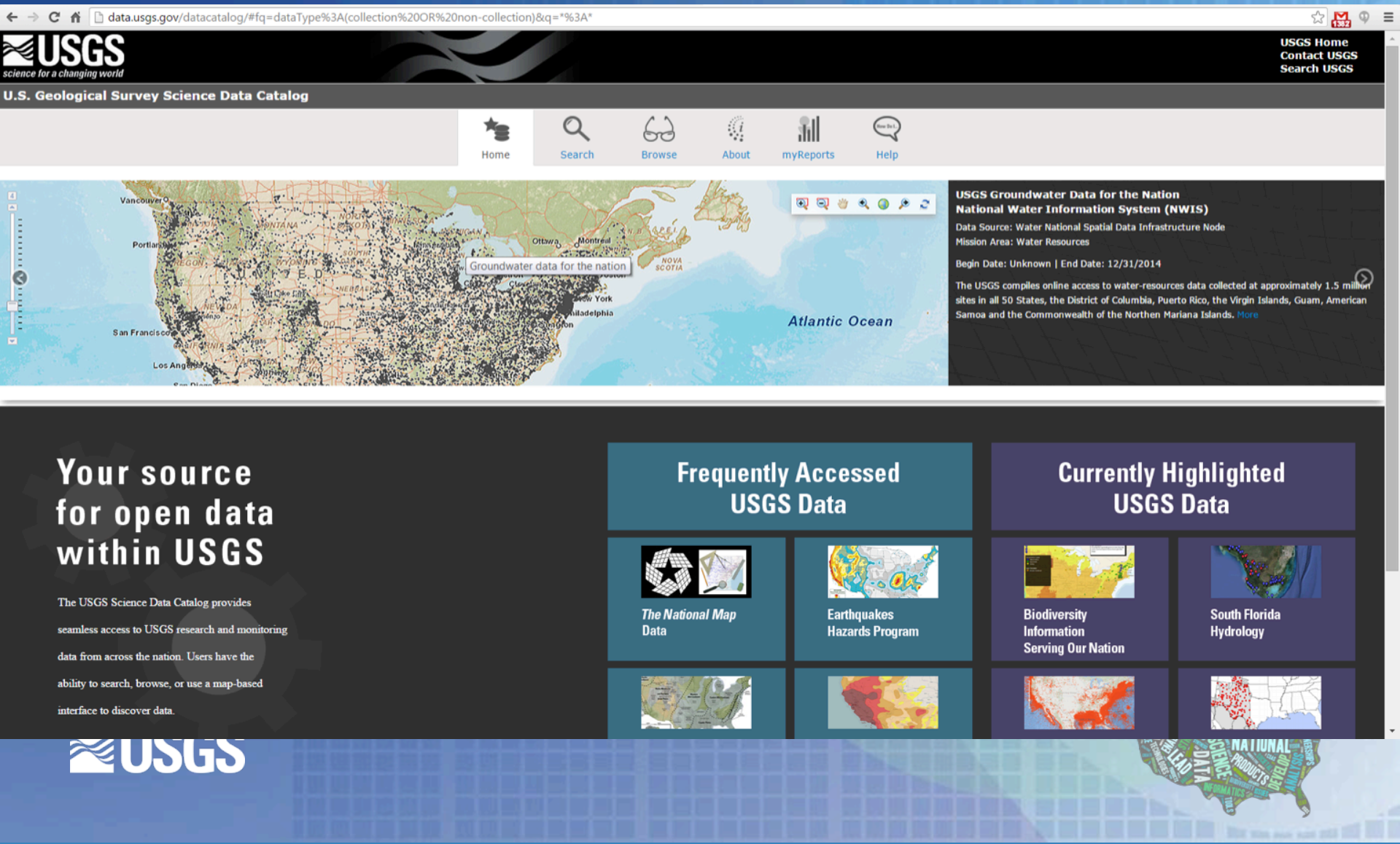
ched

2015-2016 SDC Redesign

- Goals:
  - Accommodate growth in Catalog (>doubling)
    - USGS Data Management IMs/Policies
    - Browse functionality
    - Improve highlighting of records
    - Incorporate large collections
  - Usability, look and feel
    - Testing at UT usability lab
  - Visualizing data
  - Capturing harvest history and reporting

USGS

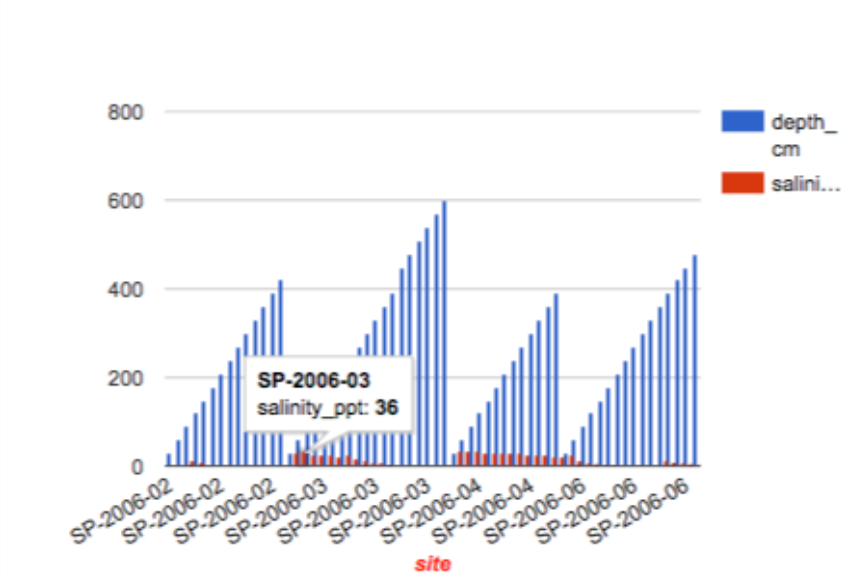
2015-2016 SDC Redesign



Visualization

Comma-delimited Text File of the Porewater Salinity Values of Cores Collected August, 2006 in the Nauset Marsh Area of Cape Cod, Massachusetts

Edit the Chart



Columns (click x to remove columns):

Add a column

x depth_cm

x salinity_ppt

USGS

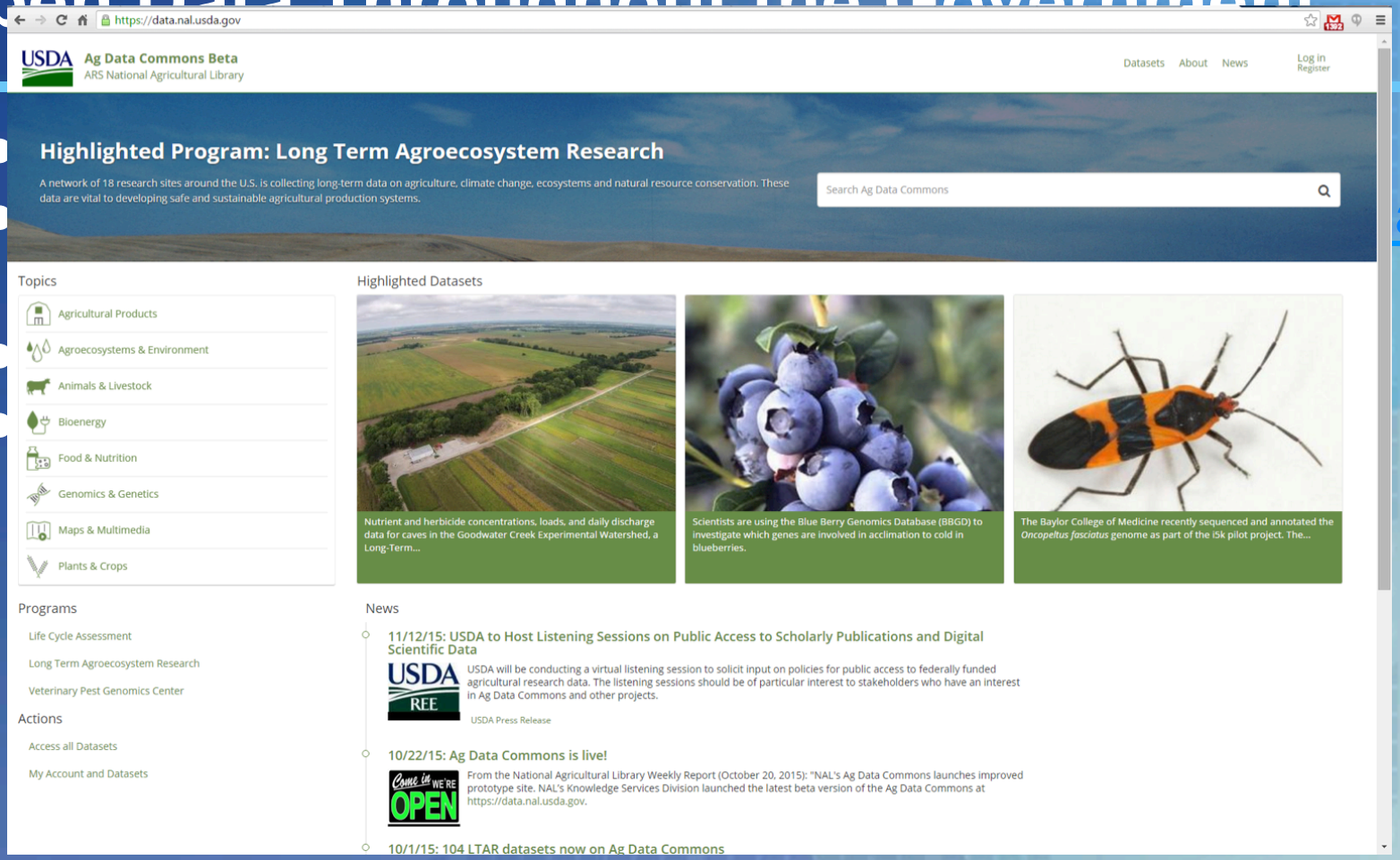
Open Data Throughout the Government

- https://data.commerce.gov/
- https://data.nasa.gov/, https://open.nasa.gov/open-data/, https://api.nasa.gov/
- https://data.nal.usda.gov/
- ……

USGS

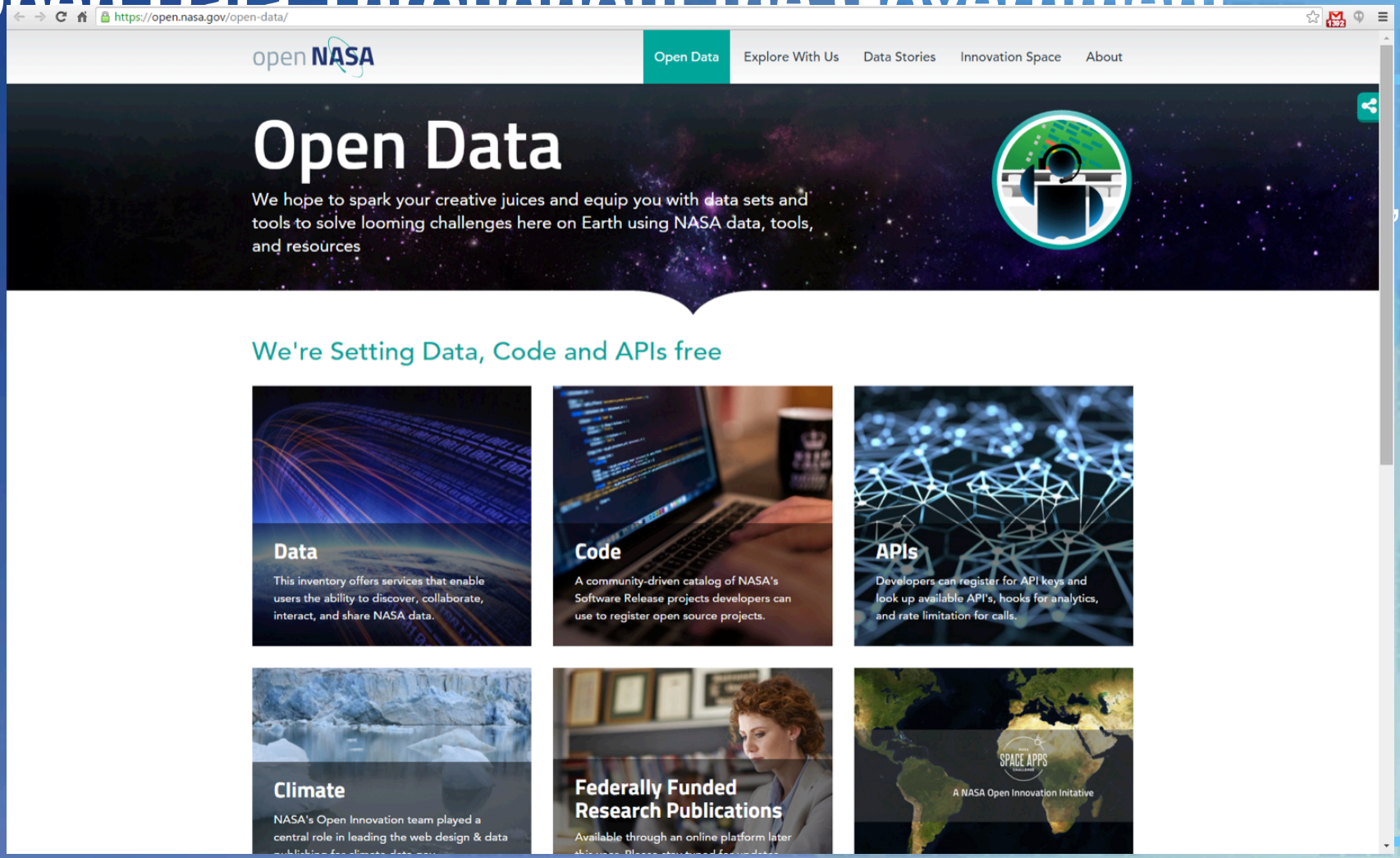
Open Data Throughout the Government

- ata/,



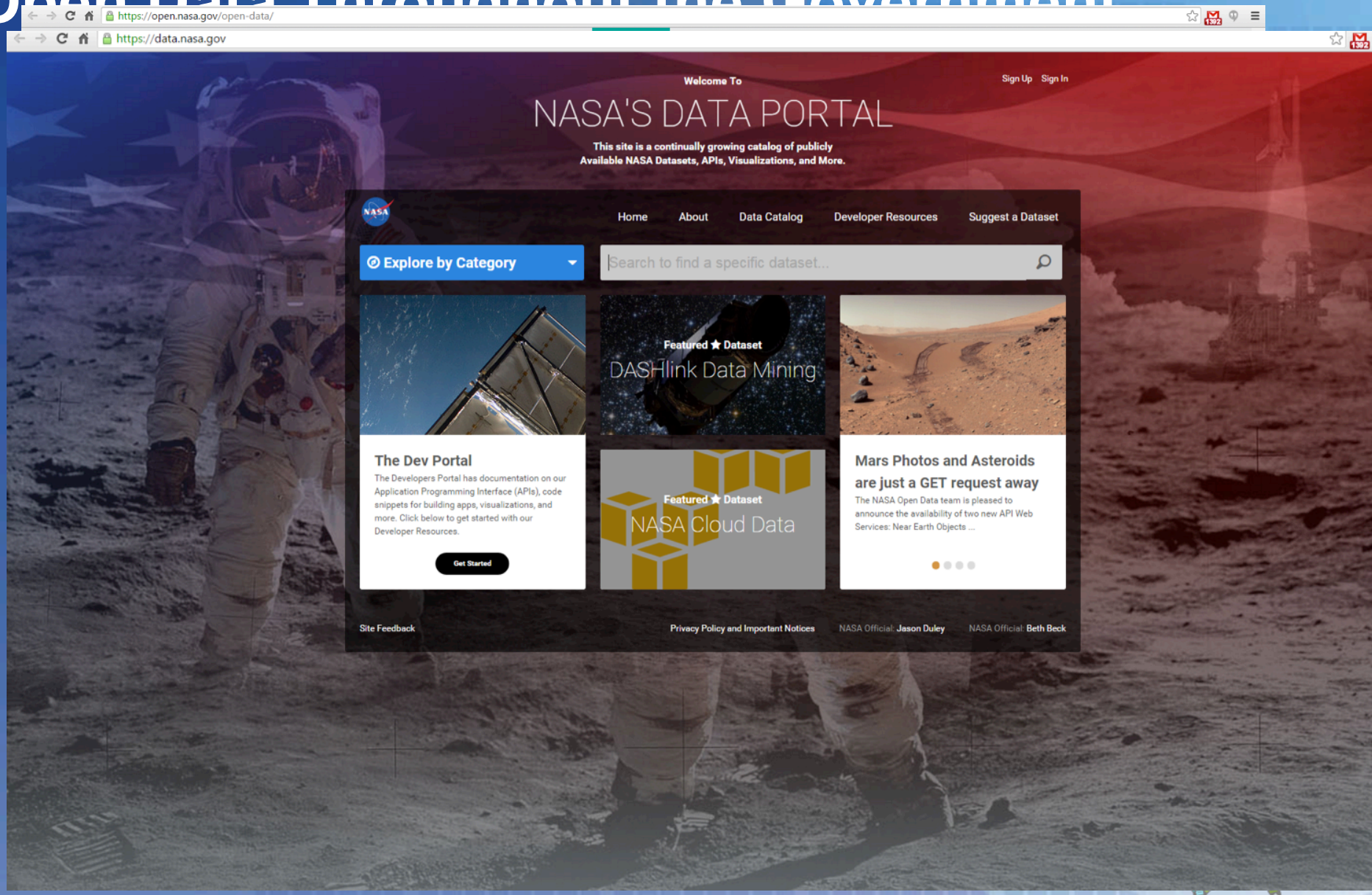
USGS

Open Data Throughout the Government



USGS

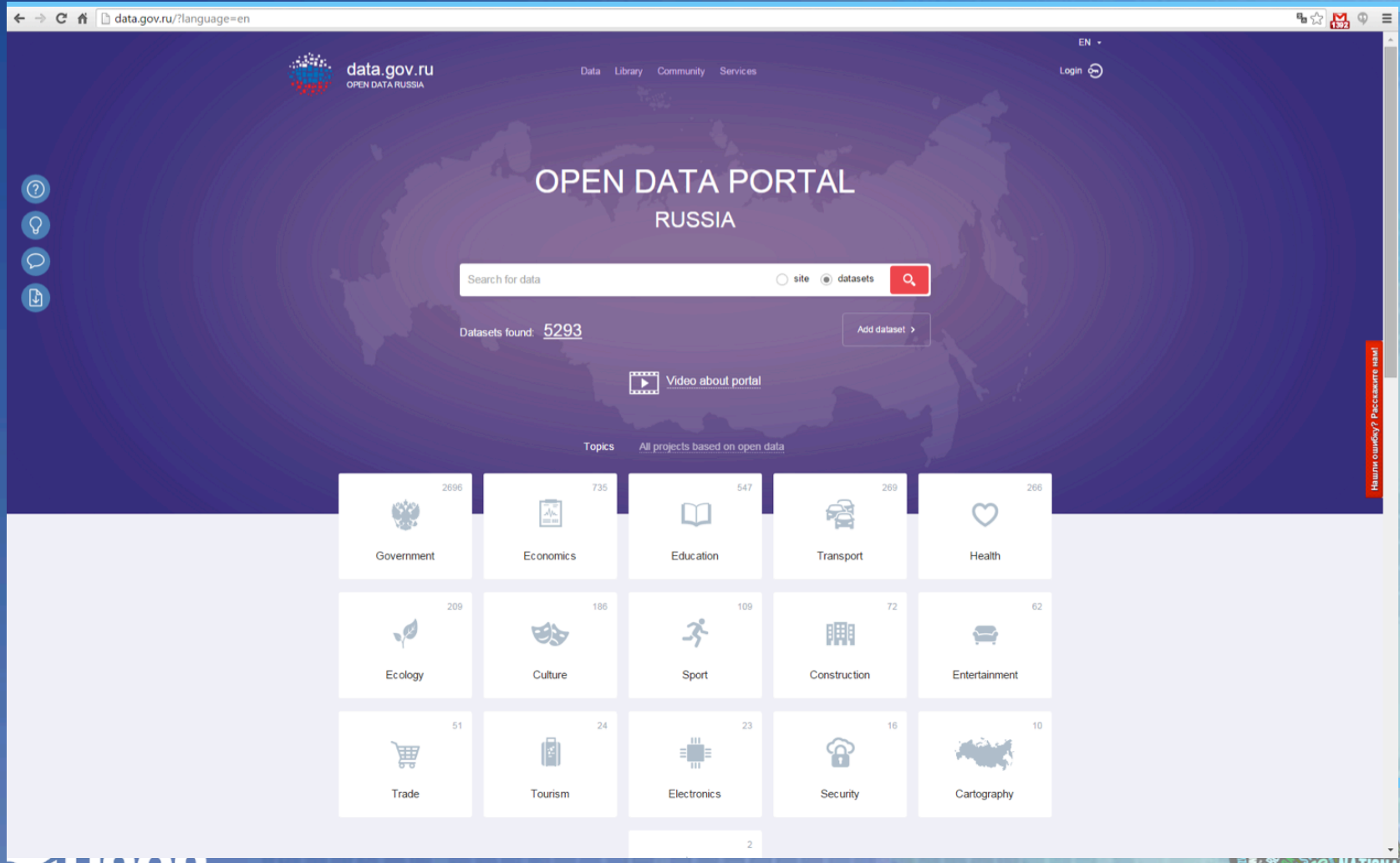
Open Data Throughout the Government



Future of Open Data

- ++ Technology
- + More and more .gov Open Data sites, growing support for concept
- +/- Increasing collaboration w/private industry
- +/- Growing power of outside interest groups, lobbyists
- ------ Funding challenges--unfunded mandates
- -/+ Institutional culture

USGS

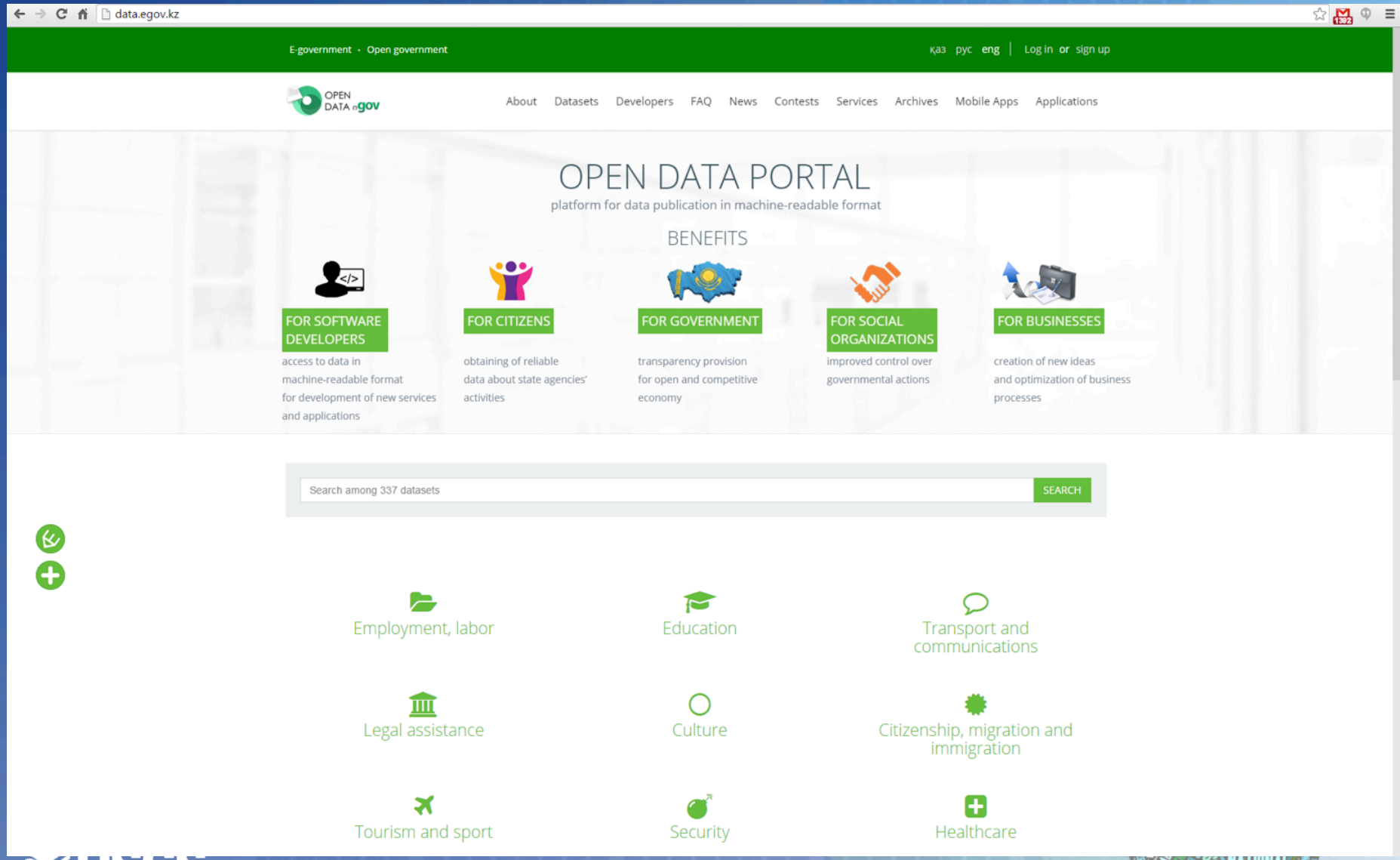

data.gov.ru/?language=en
data.gov.ru
OPEN DATA RUSSIA
Data
Library
Community
Services
EN
Login
OPEN DATA PORTAL
RUSSIA
Search for data
site
datasets
Datasets found: 5293
Add dataset
Video about portal
Topics
All projects based on open data
2696 Government
735 Economics
547 Education
269 Transport
266 Health
209 Ecology
186 Culture
109 Sport
72 Construction
62 Entertainment
51 Trade
24 Tourism
23 Electronics
16 Security
10 Cartography
2
Нашли ошибку? Расскажите нам!

USGS

NATIONAL
DATA
SCIENCE
PRODUCTS
DEVELOP
LEAD

opendata.ba/?q=tags/republika-srpska
OpenData
u Bosni i Hercegovini
search
O nama
Datasets
Groups
Log in
Register
Home / Republika Srpska
Republika Srpska
Budžet Republike Srpske
9x csv
Statistički godišnjak Republike Srpske 2014
Izaje: Republički zavod za statistiku Republike Srpske, Banja Luka, Veljka Mlađenovića 12d
1x pdf 8x csv 1x
Fondacija CPI | Powered by DKAN, a project of NuCivic

USGS


data.egov.kz
E-government · Open government
қаз рус eng | Log in or sign up
OPEN DATA gov
About Datasets Developers FAQ News Contests Services Archives Mobile Apps Applications
OPEN DATA PORTAL
platform for data publication in machine-readable format
BENEFITS
FOR SOFTWARE DEVELOPERS
access to data in machine-readable format for development of new services and applications
FOR CITIZENS
obtaining of reliable data about state agencies' activities
FOR GOVERNMENT
transparency provision for open and competitive economy
FOR SOCIAL ORGANIZATIONS
improved control over governmental actions
FOR BUSINESSES
creation of new ideas and optimization of business processes
Search among 337 datasets
SEARCH
Employment, labor
Education
Transport and communications
Legal assistance
Culture
Citizenship, migration and immigration
Tourism and sport
Security
Healthcare

Contacts

Ben Wheeler: bwheeler@usgs.gov

SDC: data.usgs.gov, sciencedatacatalog@usgs.gov
DOI CKAN: data.doi.gov
data.gov: data.gov
Geospatial Platform: geoplatform.gov

- USGS SDC Team: Mike Frame, Viv Hutchison, Lisa Zolly, Michelle Chang, Madison Langseth, Becky Uribe

USGS

We
data!